MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2018/2019

BEC2044 –ECONOMETRICS I

(All sections / Groups)

24 OCTOBER 2018
2.30 p.m. - 4.30 p.m.
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This Question paper consists of **EIGHT (8)** pages with **FOUR (4)** questions only excluding cover page.
2. Attempt **ALL** questions. The distribution of the marks for each question is given.
3. Write all your answers in the answer booklet provided.
4. Formulas and statistical tables are attached.

QUESTION 1 (40 marks)

A group of researchers from U.S. were going to do a Halloween story on the legend of full moons affecting behaviour in strange ways. They collected data from a local hospital on emergency room cases for the period from January 1, 2018 until mid-August. There were 229 observations. During this time there were eight full moons and seven new moons (a related myth concerns new moons) and three holidays (New Year's Day, Memorial Day, and Easter). If there is a full-moon effect, then hospital administrators will adjust numbers of emergency room doctors and nurses, and local police may change the number of officers on duty. Using the data collected, they obtain the regression results as shown in **Table 1**.

Table 1. Estimation Results

Dependent Variable: *Number of Emergency Room Cases*

Independent Variable	Coefficient	Std. Error
Constant	93.6958	1.5592
T	0.0338	0.0111
HOLIDAY	13.8629	6.4452
FRIDAY	6.9098	2.1113
SATURDAY	10.5894	2.1184
FULLMOON	2.4545	3.9809
NEWMOON	6.4059	4.2569

$R^2 = 0.1736$ Jarque-Bera statistic = 3.78

where *T* = a time trend (T = 1, 2, 3, ..., 229);
HOLIDAY = 1 if the day is a holiday; 0 otherwise;
FRIDAY = 1 if the day is a Friday; 0 otherwise;
SATURDAY = 1 if the day is a Saturday; 0 otherwise;
FULLMOON = 1 if there is a full moon; 0 otherwise;
NEWMOON = 1 if there is a new moon; 0 otherwise.

(a) Interpret each of the coefficients (including the constant). (7 marks)

(b) Compute the value of adjusted R^2. (2 marks)

(c) Is the error term normally distributed? Use 5% as the level of significance to test the appropriate hypothesis. (5 marks)

(d) Is the overall regression significant? Use 5% as the level of significance to test the appropriate hypothesis. (6 marks)

Continued...

(e) Develop and test (use 5% as the level of significance) appropriate hypotheses for the coefficients of *HOLIDAY*, *FRIDAY*, *SATURDAY*, *FULLMOON* and *NEWMOON*. Which coefficients are not significantly different from zero? Which variable(s) would you consider to omit from the model? (15 marks)

(f) When should emergency rooms expect more calls? (5 marks)

QUESTION 2 (20 marks)

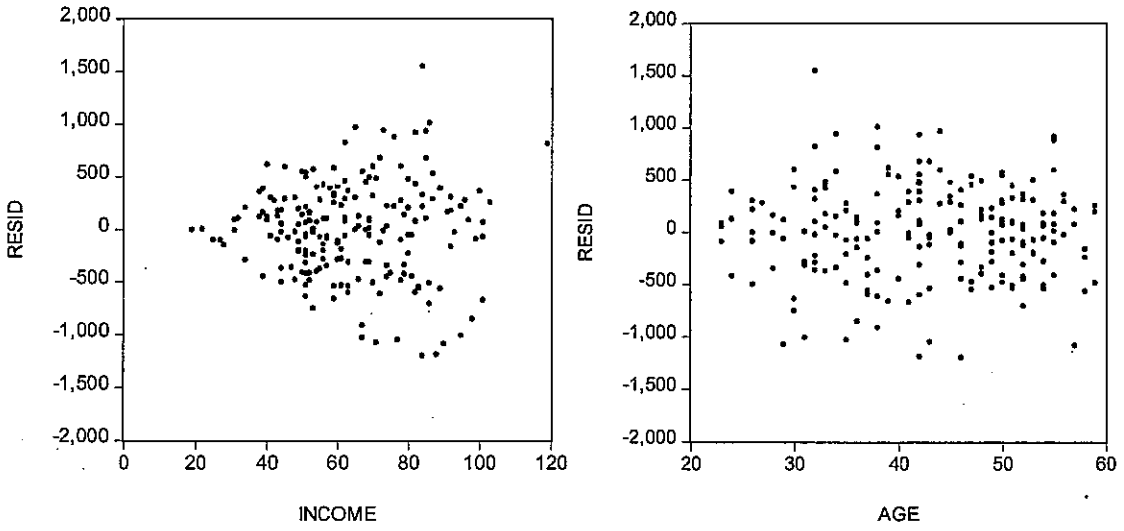
A sample of 200 Chicago households was taken to investigate how far American households tend to travel when they take vacation. Measuring distance in miles per year, the following model was estimated

$$MILES = \beta_1 + \beta_2 INCOME + \beta_3 AGE + \beta_4 KIDS + \varepsilon \quad \ldots (2.1)$$

The variables are self-explanatory except perhaps for AGE that is the average age of the adult members of the household.

(a) The equation was estimated by least squares and the residuals are plotted against age and income in **Figure 1**. What do these graphs suggest to you? (5 marks)

Figure 1. Residual Plots against Income and Age



Continued...

(b) Explain **ONE (1)** cause for **EACH** of the following econometric problems:

(i) Pure heteroscedasticity (3 marks)

(ii) Impure heteroscedasticity (3 marks)

(c) **Table 2** contains three sets of estimates, those from least squares, those from least squares with White's standard errors, and those from generalized least squares under the assumption $\sigma_t^2 = \sigma^2 \times INCOME^2$.

Table 2. Output of Three Sets of Estimates

Variable	Coefficient	Std. Error	t-Value	p-Value
Least squares estimates				
C	-391.55	169.78	-2.31	0.022
INCOME	14.20	1.80	7.89	0.000
AGE	15.74	3.76	4.19	0.000
KIDS	-81.83	27.13	-3.02	0.003
Least squares estimates with White standard errors				
C	-391.55	142.65	-2.74	0.007
INCOME	14.20	1.94	7.32	0.000
AGE	15.74	3.97	3.97	0.000
KIDS	-81.83	29.15	-2.81	0.006
Generalized least squares estimates				
C	-425.00	121.44	-3.50	0.001
INCOME	13.95	1.48	9.42	0.000
AGE	16.72	3.02	5.53	0.000
KIDS	-76.81	21.85	-3.52	0.001

(i) How do vacation miles traveled depend on income, age, and the number of kids in the household? (3 marks)

(ii) How do White's standard errors compare with the least squares standard errors? Do they change your assessment of the precision of estimation? (3 marks)

(iii) Is there evidence to suggest the generalized least squares estimates are better estimates? (3 marks)

Continued...

QUESTION 3 (20 marks)

(a) Briefly explain the following concept of specification errors and their consequences:
(i) Incorrect functional form (2 marks)
(ii) Omitted variables (2 marks)
(iii) Irrelevant variables (2 marks)

(b) Consider the following models:

Model 1 Output
Dependent variable: $\hat{Y}$

Variable	Coefficient	Std. Error
X1	1.0190	1.9989
X2	0.0073	0.0000
X3	-0.2569	0.0956
C	2.3318	9.8929
Adj R-squared	0.977	
DW d	1.98	
N	30	

Model 2 Output
Dependent variable: $\hat{Y}$

Variable	Coefficient	Std. Error
X1	-71.7243	102.8368
X3	6.6123	4.5028
C	-29.8351	522.8092
Adj R-squared	0.049	
DW d	0.23	
N	30	

(i) Using appropriate specification criteria, determine whether **Model 1** or **Model 2** is mis-specified. Explain your reasoning. (6 marks)

(ii) Test autocorrelation in both models using the Durbin-Watson d test at the 5 percent significant level. What do you conclude? (8 marks)

QUESTION 4 (20 marks)

(a) A researcher wishes to investigate the relationship between unemployment rate and job vacancy rate. Since he is using quarterly data, he is concerned with seasonal effects on the unemployment rate. Thus, he includes 3 dummy variables to capture the seasonal effects. Below is the estimation result, with the standard errors (s.e.) in parentheses:

$$\widehat{UNEMP}_t = \underset{(s.e.)}{3.2982} - \underset{(0.2567)}{1.9839VAC_t} - \underset{(0.0009)}{0.0030DQ1_t} - \underset{(0.157)}{0.0477DQ2_t} - \underset{(0.0250)}{0.0625DQ3_t} \quad \ldots(4.1)$$

where $UNEMP$ = unemployment rate (%);
VAC = job vacancy rate (%);
$DQ1$ = 1 if quarter 1 and 0 otherwise;
$DQ2$ = 1 if quarter 2 and 0 otherwise;
$DQ3$ = 1 if quarter 3 and 0 otherwise.

Continued...

(i) Compute the average unemployment rate for each quarter. (6 marks)

(ii) Explain why only 3 dummy variables are included in your regression model. (2 marks)

(b) You wish to investigate the money supply model of Malaysia using the following equation:

$$M2_t = \alpha_0 + \alpha_1 GDP_t + \alpha_2 CPI_t + \alpha_3 IntR_t + u_t \quad \ldots(4.2)$$

where $M2$ = money supply (in RM)
GDP = gross domestic product (in RM)
CPI = consumer price index
$IntR$ = interest rate (in percentage)
u = error term

Prior to perform the regression, you obtain the correlation matrix of the variables as below:

	M2	**GDP**	**CPI**	**IntR**
M2	1.000	0.992	0.897	-0.456
GDP	0.992	1.000	0.940	-0.421
CPI	0.897	0.940	1.000	-0.267
IntR	-0.456	-0.421	-0.267	1.000

(i) From the correlation matrix output, can you detect any econometric problems? Explain. (4 marks)

(ii) Given your answer in (i), what are the consequential impacts on the properties of the least squares estimators and the estimated standard errors if there is an econometric problem detected in the model? Briefly explain. (4 marks)

(iii) Suggest **TWO (2)** remedies that you may use to solve the problem you answered in (i). Briefly explain. (4 marks)

End of Page

Formulas

$$\text{t - statistic} = \frac{\hat{\beta} - \beta}{\text{se}(\hat{\beta})}$$

$$\text{F - statistic} = \frac{\text{ESS}/(k)}{\text{RSS}/(n-k-1)} = \frac{R^2/(k)}{(1-R^2)/(n-k-1)}$$

$$R^2 = \frac{\text{ESS}}{\text{TSS}}$$

$$\text{Adjusted } R^2 = 1 - (1 - R^2)\left(\frac{n-1}{n-k-1}\right)$$

Total Sum of Squares (TSS) = Explained Sum of Squares (ESS) + Residual Sum of Squares (RSS)

White's general heteroscedasticity test statistics, $\chi^2_{df=k} = n \times R^2$

$\rho \approx 1 - (d/2)$

Statistical Tables

Appendix 1: t-Table

two tails	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
One tail	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**
df					
10	1.37	1.81	2.23	2.76	3.17
20	1.33	1.72	2.09	2.53	2.84
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
50	1.30	1.68	2.01	2.40	2.68
60	1.30	1.67	2.00	2.39	2.66
75	1.29	1.67	1.99	2.38	2.64
100	1.29	1.66	1.98	2.36	2.63
120	1.29	1.66	1.98	2.36	2.62

Appendix 2: F-table (α=0.05)

df2\df1	1	2	3	4	5	6	7	8
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06
100	3.94	3.09	2.7	2.46	2.31	2.19	2.10	2.03
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96
1000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95

Appendix 3: Chi-square table (α=0.05)

df	α = 0.05
1	3.84
2	5.99
3	7.82
4	9.49
5	11.07
10	18.31
20	31.41
30	43.77
40	55.76
50	67.51
100	124.34

Appendix 4: Durbin Watson d Table

Critical Values for the Durbin-Watson Statistic (d)

Level of Significance $\alpha = 0.05$

n	$k=1$		$k=2$		$k=3$		$k=4$		$k=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
10	0.88	1.32	0.70	1.64	0.53	2.02	0.38	2.41	0.24	2.82
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
25	1.29	1.41	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78
150	1.72	1.75	1.71	1.76	1.69	1.77	1.68	1.79	1.66	1.80
200	1.76	1.78	1.75	1.79	1.74	1.80	1.73	1.81	1.72	1.82

n	$k=6$		$k=7$		$k=8$		$k=9$		$k=10$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
10	–	–	–	–	–	–	–	–	–	–
15	0.45	2.47	0.34	2.73	0.25	2.98	0.18	3.22	0.11	3.44
20	0.69	2.16	0.60	2.34	0.50	2.52	0.42	2.70	0.34	2.89
25	0.87	2.01	0.78	2.14	0.70	2.28	0.62	2.42	0.54	2.56
30	1.00	1.93	0.93	2.03	0.85	2.14	0.78	2.25	0.71	2.36
40	1.18	1.85	1.12	1.92	1.06	2.00	1.01	2.07	0.95	2.15
50	1.29	1.82	1.25	1.88	1.20	1.93	1.16	1.99	1.11	2.04
60	1.37	1.81	1.34	1.85	1.30	1.89	1.26	1.94	1.22	1.98
70	1.43	1.80	1.40	1.84	1.37	1.87	1.34	1.91	1.31	1.95
80	1.48	1.80	1.45	1.83	1.43	1.86	1.40	1.89	1.37	1.93
90	1.52	1.80	1.49	1.83	1.47	1.85	1.45	1.88	1.42	1.91
100	1.55	1.80	1.53	1.83	1.50	1.85	1.48	1.87	1.46	1.90
150	1.65	1.82	1.64	1.83	1.62	1.85	1.60	1.86	1.59	1.88
200	1.71	1.83	1.70	1.84	1.69	1.85	1.68	1.86	1.67	1.87

where n = number of observations and k = number of independent variables, excluding the intercept.